

Uplink Performance of Conventional and Massive MIMO Cellular Systems with Delayed CSIT

Anastasios K. Papazafeiropoulos, Hien Quoc Ngo, and Tharm Ratnarajah

Abstract

Massive multiple-input multiple-output (MIMO) networks, where the base stations (BSs) are equipped with large number of antennas and serve a number of users simultaneously, are very promising, but suffer from pilot contamination. Despite its importance, delayed channel state information (CSI) due to user mobility, being another degrading factor, lacks investigation in the literature. Hence, we consider an uplink model, where each BS applies zero-forcing decoder, accounting for both effects, but with the focal point on the relative users' movement with regard to the BS antennas. In this setting, analytical closed-form expressions for the sum-rate with finite number of BS antennas, and the asymptotic limits with infinite number of BS antennas epitomize the main contributions. In particular, the probability density function of the signal-to-interference-plus-noise ratio and the ergodic sum-rate are derived for any finite number of antennas. Insights of the impact of the arising Doppler shift due to user mobility into the low signal-to-noise ratio regime as well as the outage probability are obtained. Moreover, asymptotic analysis performance results in terms of infinitely increasing number of antennas, power,

Parts of this work were presented at the 2014 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) [1].

A. K. Papazafeiropoulos is with Communications and Signal Processing Group, Imperial College London, London, U.K. (email: a.papazafeiropoulos@imperial.ac.uk).

H. Q. Ngo is with the Department of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden (email: nqhien@isy.liu.se).

T. Ratnarajah is with Institute for Digital Communications (IDCoM), University of Edinburgh, Edinburgh, U.K. (email: t.ratnarajah@ed.ac.uk).

This research was supported by a Marie Curie Intra European Fellowship and HARP project within the 7th European Community Framework Programme for Research of the European Commission under grant agreements no. [330806], IAWICOM and no. [318489], HARP.

and both numbers of antennas and users (while their ratio is fixed) are provided. The numerical results demonstrate the performance loss in various Doppler shifts. An interesting observation is that massive MIMO is favorable even in time-varying channel conditions.

Index Terms

Delayed channels, multiuser multiple-input multiple-output (MIMO), massive MIMO, zero-forcing (ZF).

I. INTRODUCTION

The rapidly increasing demand for wireless connectivity and throughput is one of the motivations for the continuous evolution of cellular networks [2], [3]. Very large multiple-input multiple-output (MIMO) has been identified as a new promising breakthrough technique aiming at achieving higher area throughput in wireless networks [4]–[6]. Its origin is found in [4], and it has been given many alternative names such as massive MIMO, hyper-MIMO, and full-dimension MIMO systems. In the typical envisioned architecture, each base station (BS) with an array of hundreds or even thousands antennas, exploiting the key idea of multi-user MIMO (MU-MIMO), serves tens or hundreds of single-antenna users simultaneously in the same frequency band, respectively, under coherent processing. This difference in the number of BS antennas N and the number of users K per cell provides unprecedented spatial degrees of freedom that leads to high signal gains, allowing at the same time low-complexity linear signal processing techniques and avoiding inter-user interference due to the (near) orthogonality between the channels.

On a similar note, zero-forcing (ZF) processing is regarded as a low-complexity alternative of maximum-likelihood multiuser detector and “dirty paper coding” [7], especially, when the BSs are equipped with massive antenna arrays. A lot of research has been conducted on single-cell systems with ZF receivers [8], but the main current interest has shifted to practical multi-cell scenarios, where pilot contamination degrades the system performance [4], [9].

Despite that the theory of massive MIMO has been now well established (see [4] and references therein), an important question that has been overlooked is how the performance of massive MIMO topology is affected by the relative movement of users. This scenario is of high practical importance, in urban environments, where users move rapidly within a geographical area. The main challenge in time-varying environments is to perform robust channel estimation, when the

propagation channel changes over time. The dynamic channel behavior was modeled in terms of a stationary ergodic Gauss-Markov block fading channel model [10]–[12], where an autoregressive model was combined with the Jakes’ autocorrelation function that captures the time variation of the channel. Motivated by the above observation, this paper explores the robustness offered by massive MIMO against the practical setting of user mobility that results to delayed and degraded channel state information (CSI) at the BS, and thus, imperfect CSI. Such consideration is notably important because it can provide the quantification of the performance loss in various Doppler shifts.

A limited effort for studying the time variation of the channel due to the relative movement of users has been conducted in [10], where the authors provided deterministic equivalents (DEs)¹ for the maximal ratio combining (MRC) receiver in the uplink and the matched filter (MF) in the downlink transmission. Fortunately, this analysis was extended in [11], [12] by deriving DEs for the minimum mean-square error (MMSE) receiver (uplink) and regularized zero-forcing (downlink) and by making a comparison regarding their performance. In this paper, we consider a generalized uplink massive MIMO system. Based on the aforementioned literature, we propose a tractable model that encompasses ZF receivers and describes the impact of user mobility in a cellular system with BSs having conventional and very large number of antennas, or even large number of both antennas and users, which stands in contrast to the previous works. The following are the main contributions of this paper:

- Contrary to [15], we consider more practical settings where the channel is imperfectly estimated at the BS. The effects of pilot contamination and time variation of the channels are taken in to account. The extension is not straightforward because apart of the development of the model, the mathematical manipulations are hampered. Apart of this, the results are contributory and novel.
- We derive the probability density function (PDF) of the signal-to-interference-plus-noise ratio (SINR) and the corresponding ergodic sum-rate for any finite number of antennas in closed forms. For the sake of completeness, the link of these results with previous

¹The deterministic equivalents are deterministic tight approximations of functionals of random matrices of finite size. Note that these approximations are asymptotically accurate as the matrix dimensions grow to infinity, but can be precise for small dimensions.

known results is mentioned. Furthermore, a simpler and more tractable lower bound for the achievable uplink rate is derived.

- We elaborate on the low signal-to-noise ratio (SNR) regime, in order to get additional insights into the impact of Doppler shift. In particular, we study the behaviors of the minimum normalized energy per information bit to reliably convey any positive rate and the wideband slope.
- We present a simple expression for the outage probability, being an important metric in quasi-static models.
- We investigate the asymptotic performance presented by very large MIMO ($N \rightarrow \infty$) as well as large MIMO in terms of DEs ($N, K \rightarrow \infty$). This analysis aims at providing accurate approximation results that replace the need for lengthy Monte Carlo simulations.

Note that, although all the results incur significant mathematical challenges, they can be easily evaluated. Nevertheless, the purpose of DEs is to provide the deterministic tight approximations, in order to avoid lengthy Monte-Carlo simulations.

Notation: We use boldface lowercase and uppercase letters to denote vectors and matrices, respectively. The notation $(\cdot)^H$ stands for the conjugate transpose, and $\|\cdot\|$ denotes the Euclidean norm of a vector, while $(\cdot)^\dagger$ denotes the pseudo-inverse of a matrix. We use the notation $x \stackrel{d}{\sim} y$ to imply that x and y have the same distribution. Finally, we use $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ to denote a circularly symmetric complex Gaussian vector \mathbf{z} with zero mean and covariance matrix $\mathbf{\Sigma}$.

II. SYSTEM MODEL

Consider a cellular network which has L cells. Each cell includes one N -antenna BS and K single-antenna users. We consider the uplink transmission. The model is based on the assumptions that: i) $N \geq K$, and ii) all users in L cells share the same time-frequency resource. Moreover, we assume that the channels are frequency flat and they vary from symbol to symbol, while during the symbol period they are considered constant due to the channel aging impact [10] (we will discuss about the channel aging model later). The channel vector $\mathbf{g}_{lik}[n] \in \mathbb{C}^{N \times 1}$ between the k th user in the i th cell and the l th BS at the n th symbol undergoes independent small-scale fading and large-scale fading. More precisely, $\mathbf{g}_{lik}[n]$ is modelled as

$$\mathbf{g}_{lik}[n] = \sqrt{\beta_{lik}} \mathbf{h}_{lik}[n], \quad (1)$$

where β_{lik} represents large-scale fading, and $\mathbf{h}_{lik} \in \mathbb{C}^{N \times 1} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ is the small-scale fading vector between the l th BS and the k th user in the i th cell.

Let $\sqrt{p_r}\mathbf{x}_i[n] \in \mathbb{C}^{K \times 1}$ be the zero-mean stochastic data signal vector of K users in the i th cell at time instance n ($p_r > 0$ is the average transmitted power of each user, and $\mathbf{x}_i[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$). Then, the $N \times 1$ received signal vector at the l th BS is

$$\mathbf{y}_l[n] = \sqrt{p_r} \sum_{i=1}^L \mathbf{G}_{li}[n] \mathbf{x}_i[n] + \mathbf{z}_l[n], \quad l = 1, 2, \dots, L, \quad (2)$$

where $\mathbf{G}_{li}[n] \triangleq [\mathbf{g}_{li1}[n], \dots, \mathbf{g}_{liK}[n]] \in \mathbb{C}^{N \times K}$ denotes the channel matrix between the K users in the i th cell and the l th BS, and $\mathbf{z}_l[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ is additive white Gaussian noise (AWGN) vector at the l th BS.

To coherently detect the signals transmitted from the K users in the l th cell, the BS needs CSI knowledge. Conventionally, the l th BS can estimate the channel via uplink training. We assume that the channel remains constant during the training phase [10]. In general, this assumption is not practical, but it yields a simple model which enables us to analyze the system performance and to obtain initial insights on the impact of channel aging. Furthermore, the impact of channel aging can be absorbed in the channel estimation error.

During the training phase, in each cell, K users are assigned K orthogonal pilot sequences of length τ symbols ($\tau \geq K$). Owing to the limitation of the coherence interval, the pilot sequences have to be reused from cell to cell. We assume that all L cells use the same set of orthogonal pilot sequences. As a result, the pilot contamination occurs [4], [5]. Denote by $\mathbf{\Psi} \in \mathbb{C}^{K \times \tau}$, ($\tau \geq K$), be the pilot matrix transmitted from the K users in each cell, where the k th row of $\mathbf{\Psi}$ is the pilot sequence assigned for the k th user. The matrix $\mathbf{\Psi}$ satisfies $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{I}_K$. Then, the $N \times \tau$ received pilot signal at BS l is

$$\mathbf{Y}_l^{\text{tr}}[n] = \sqrt{p_{\text{tr}}} \sum_{i=1}^L \mathbf{G}_{li}[n] \mathbf{\Psi} + \mathbf{Z}_l^{\text{tr}}[n], \quad l = 1, 2, \dots, L, \quad (3)$$

where the subscript $()^{\text{tr}}$ implies the uplink training stage, $p_{\text{tr}} \triangleq \tau p_r$, and $\mathbf{Z}_l^{\text{tr}}[n] \in \mathbb{C}^{N \times \tau}$ is spatially AWGN at BS l during the training phase. We assume the elements of $\mathbf{Z}_l^{\text{tr}}[n]$ are i.i.d. $\mathcal{CN}(0, 1)$ random variables (RVs). The MMSE channel estimate of $\mathbf{g}_{lik}[n]$ is given by [10]

$$\hat{\mathbf{g}}_{lik}[n] = \beta_{lik} \mathbf{Q}_{lk} \left(\sum_{j=1}^L \mathbf{g}_{ljk}[n] + \frac{1}{\sqrt{p_{\text{tr}}}} \tilde{\mathbf{z}}_{lk}^{\text{tr}}[n] \right), \quad (4)$$

where $\mathbf{Q}_{lk} \triangleq \left(\frac{1}{p_{\text{tr}}} + \sum_{i=1}^L \beta_{lik} \right)^{-1} \mathbf{I}_N$, and $\tilde{\mathbf{z}}_{lk}^{\text{tr}}[n] \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ represents the additive noise.

With MMSE channel estimation, the channel estimate and the channel estimation error are uncorrelated. Thus, $\mathbf{g}_{lik}[n]$ can be rewritten as:

$$\mathbf{g}_{lik}[n] = \hat{\mathbf{g}}_{lik}[n] + \tilde{\mathbf{g}}_{lik}[n], \quad (5)$$

where $\tilde{\mathbf{g}}_{lik}[n] \sim \mathcal{CN}(\mathbf{0}, (\beta_{lik} - \hat{\beta}_{lik}) \mathbf{I}_N)$ and $\hat{\mathbf{g}}_{lik}[n] \sim \mathcal{CN}(\mathbf{0}, \hat{\beta}_{lik})$, with $\hat{\beta}_{lik} = \frac{\beta_{lik}^2}{\sum_{j=1}^L \beta_{ljk} + 1/p_{\text{tr}}}$, are the independent channel estimate and channel estimation error, respectively. Note that β_{lik} , $\hat{\beta}_{lik}$, and \mathbf{Q}_{lk} are independent of $n \forall l, i$, and k .

Besides pilot contamination, in any common propagation scenario, a relative movement takes place between the antennas and the scatterers that degrades more channel's performance. Under these circumstances, the channel is time-varying and needs to be modeled by the famous Gauss-Markov block fading model, which is basically an autoregressive model of certain order that incorporate two-dimensional isotropic scattering (Jakes model). More specifically, our analysis achieves to relate the current channel state with its past samples. For the sake of analytical simplicity, we consider the following simplified autoregressive model of order 1 [10]

$$\mathbf{g}_{lik}[n] = \alpha \mathbf{g}_{lik}[n-1] + \mathbf{e}_{lik}[n], \quad (6)$$

where $\mathbf{g}_{lik}[n-1]$ and $\mathbf{e}_{lik}[n] \sim \mathcal{CN}(\mathbf{0}, (1 - \alpha^2) \beta_{lik} \mathbf{I}_N)$ are uncorrelated, denoting the channel at the previous symbol duration and the stationary Gaussian channel error vector due to the time variation of the channel, respectively. Note that $\alpha = J_0(2\pi f_D T_s)$, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_D and T_s are the maximum Doppler shift and the channel sampling period. Basically, α , which is assumed known at the BS, corresponds to the temporal correlation parameter that describes the isotropic scattering according to the Jakes' model. In particular, the maximum Doppler shift f_D equals $f_D = \frac{vf_c}{c}$, where v (in m/s) is the relative velocity of the user, $c = 3 \times 10^8$ m/s is the speed of light, and f_c is the carrier frequency.

Substituting (5) into (6), we obtain a model which combines both effects of channel estimation error and channel aging as follows:

$$\begin{aligned} \mathbf{g}_{lik}[n] &= \alpha \mathbf{g}_{lik}[n-1] + \mathbf{e}_{lik}[n] \\ &= \alpha \hat{\mathbf{g}}_{lik}[n-1] + \tilde{\mathbf{e}}_{lik}[n], \end{aligned} \quad (7)$$

where $\hat{\mathbf{g}}_{lik}[n-1]$ and $\tilde{\mathbf{e}}_{lik}[n] \triangleq \alpha \tilde{\mathbf{g}}_{lik}[n-1] + \mathbf{e}_{lik}[n] \sim \mathcal{CN}(\mathbf{0}, (\beta_{lik} - \alpha^2 \hat{\beta}_{lik}) \mathbf{I}_N)$ are mutually independent. More concretely, we define $\hat{\mathbf{G}}_i[n] \triangleq [\hat{\mathbf{g}}_{li1}[n], \dots, \hat{\mathbf{g}}_{liK}[n]] \in \mathbb{C}^{N \times K}$ and $\tilde{\mathbf{E}}_{li} \triangleq$

$$\gamma_k = \frac{\alpha^2 p_r}{\alpha^2 p_r \sum_{i \neq l}^L \left\| \left[\hat{\mathbf{G}}_l^\dagger[n-1] \right]_k \hat{\mathbf{G}}_{li}[n-1] \right\|^2 + p_r \sum_{i=l}^L \left\| \left[\hat{\mathbf{G}}_l^\dagger[n-1] \right]_k \tilde{\mathbf{E}}_{li}[n] \right\|^2 + \left\| \left[\hat{\mathbf{G}}_l^\dagger[n-1] \right]_k \right\|^2}. \quad (12)$$

$[\tilde{\mathbf{e}}_{li1}[n], \dots, \tilde{\mathbf{e}}_{liK}[n]] \in \mathbb{C}^{N \times K}$ as the combined channel matrices from all users in cell i to BS l . In particular, $\hat{\mathbf{G}}_{li}[n]$ can be expressed as [15]:

$$\hat{\mathbf{G}}_{li}[n] = \hat{\mathbf{G}}_l[n] \mathbf{D}_i, \quad (8)$$

where $\mathbf{D}_i = \text{diag}\left\{\frac{\beta_{li1}}{\beta_{li1}}, \frac{\beta_{li2}}{\beta_{li2}}, \dots, \frac{\beta_{liK}}{\beta_{liK}}\right\}$.

Making use of (7), we can rewrite the received signal $\mathbf{y}_l[n]$ at the l th BS ($l \in [1, L]$) as

$$\mathbf{y}_l[n] = \alpha \sqrt{p_r} \sum_{i=1}^L \hat{\mathbf{G}}_{li}[n-1] \mathbf{x}_i[n] + \sqrt{p_r} \sum_{i=1}^L \tilde{\mathbf{E}}_{li}[n] \mathbf{x}_i[n] + \mathbf{z}_l[n]. \quad (9)$$

Moreover, we assume that the l th BS uses the ZF technique to detect the signals transmitted from K users in its cells. With ZF, the received signal vector $\mathbf{y}_l[n]$ is pre-multiplied with $\alpha^{-1} \hat{\mathbf{G}}_l^\dagger[n-1]$, where $\hat{\mathbf{G}}_l^\dagger[n-1]$ is a $K \times N$ matrix representing the pseudo-inverse of $\hat{\mathbf{G}}_l[n-1]$:

$$\begin{aligned} \mathbf{r}_l[n] &= \sqrt{p_r} \mathbf{x}_l[n] + \sqrt{p_r} \sum_{i \neq l}^L \hat{\mathbf{G}}_l^\dagger[n-1] \hat{\mathbf{G}}_{li}[n-1] \mathbf{x}_i[n] \\ &+ \alpha^{-1} \sqrt{p_r} \sum_{i=1}^L \hat{\mathbf{G}}_l^\dagger[n-1] \tilde{\mathbf{E}}_{li}[n] \mathbf{x}_i[n] + \alpha^{-1} \hat{\mathbf{G}}_l^\dagger[n-1] \mathbf{z}_l[n], \end{aligned} \quad (10)$$

Then, the k th element of $\mathbf{r}_l[n]$ is used to detect the signal transmitted from the k th user. The post-processed received signal corresponding to the k th user is

$$\begin{aligned} r_{lk}[n] &= \sqrt{p_r} x_{lk}[n] + \sqrt{p_r} \sum_{i \neq l}^L \left[\hat{\mathbf{G}}_l^\dagger[n-1] \right]_k \hat{\mathbf{G}}_{li}[n-1] \mathbf{x}_i[n] \\ &+ \frac{1}{\alpha} \sqrt{p_r} \sum_{i=1}^L \left[\hat{\mathbf{G}}_l^\dagger[n-1] \right]_k \tilde{\mathbf{E}}_{li}[n] \mathbf{x}_i[n] + \frac{1}{\alpha} \left[\hat{\mathbf{G}}_l^\dagger[n-1] \right]_k \mathbf{z}_l[n], \end{aligned} \quad (11)$$

where the notation $[\mathbf{A}]_k$ refers to the k th row of matrix \mathbf{A} , and $x_{lk}[n]$ is the k th element of $\mathbf{x}_l[n]$, i.e., it is the transmit signal from the k th user in the l th cell at the n th time slot. Treating (11) as a single-input single-output (SISO) system, we obtain the SINR of the transmission from the k th user in the l th cell to its BS in (12) shown at the top of the page. The SINR is obtained under the assumption that the l th BS knows the denominator value of (12). This assumption is reasonable since this value is just a scalar (which can be estimated).

III. ACHIEVABLE UPLINK SUM RATE

In this section, we provide the sum-rate analysis for finite and infinite number of BS antennas taking into account the aforementioned effects.

A. Finite- N Analysis

Proposition 1: The uplink SINR of transmission between the k th user in the l th cell to its BS, under the delayed channels, is distributed as

$$\gamma_k \stackrel{d}{\sim} \frac{\alpha^2 p_r X_k[n-1]}{\alpha^2 p_r C X_k[n-1] + p_r Y_k[n] + 1}, \quad (13)$$

where X_k and Y_k are independent random variables (RVs) whose PDFs are, respectively, given by

$$p_{X_k}(x) = \frac{e^{-x/\hat{\beta}_{lk}}}{(N-K)! \hat{\beta}_{lk}} \left(\frac{x}{\hat{\beta}_{lk}} \right)^{N-K}, \quad x \geq 0 \quad (14)$$

$$p_{Y_k}(y) = \sum_{p=1}^{\varrho(\mathcal{A}_k) \tau_p(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \mathcal{X}_{p,q}(\mathcal{A}_k) \frac{\mu_{k,p}^{-q}}{(q-1)!} y^{q-1} e^{-\frac{y}{\mu_{k,p}}}, \quad y \geq 0, \quad (15)$$

where $\mathcal{A}_k \triangleq \text{diag}(\tilde{\mathbf{D}}_{l1}, \dots, \tilde{\mathbf{D}}_{lL}) \in \mathbb{C}^{KL \times KL}$ with $\tilde{\mathbf{D}}_{li}$ a $K \times K$ diagonal matrix having elements $[\tilde{\mathbf{D}}_{li}]_{kk} = (\beta_{lik} - \alpha^2 \hat{\beta}_{lik})$, as well as $\varrho(\mathcal{A}_k)$ denotes the numbers of distinct diagonal elements of \mathcal{A}_k . Similarly, $\mu_{k,1}, \mu_{k,2}, \dots, \mu_{k,\varrho(\mathcal{A}_k)}$ are the associated distinct diagonal elements in decreasing order and $\tau_p(\mathcal{A}_k)$ are the multiplicities of $\mu_{k,p}$, while $\mathcal{X}_{p,q}(\mathcal{A}_k)$ is the (p, q) th characteristic coefficients of \mathcal{A}_k , as defined in [16, Definition 4]. Regarding C , it is a deterministic constant: $C \triangleq \sum_{i \neq l}^L \left(\frac{\beta_{lik}}{\beta_{lk}} \right)^2$.

Proof: Division of each term of (12) by $\left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^2$ leads to

$$\gamma_k = \frac{\alpha^2 p_r \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^{-2}}{\alpha^2 p_r C \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^{-2} + p_r \sum_{i=1}^L \left\| \hat{\mathbf{Y}}_i[n] \right\|^2 + 1}, \quad (16)$$

where $C \triangleq \sum_{i \neq l}^L \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \hat{\mathbf{G}}_{li}[n-1] \right\| \stackrel{(8)}{=} \sum_{i \neq l}^L \left(\frac{\beta_{lik}}{\beta_{lk}} \right)^2$ and $\hat{\mathbf{Y}}_i[n] \triangleq \frac{[\hat{\mathbf{G}}_u^\dagger[n-1]]_k \tilde{\mathbf{E}}_{li}[n]}{\left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|}$.

Since $\left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^2 = \left[\left(\hat{\mathbf{G}}_u^H[n-1] \hat{\mathbf{G}}_u[n-1] \right)^{-1} \right]_{kk}$, $\left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^{-2}$ has an Erlang distribution with shape parameter $N - K + 1$ and scale parameter $\hat{\beta}_{lk}$ [17]. Then,

$$\left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^{-2} \stackrel{d}{\sim} X_k[n-1]. \quad (17)$$

$$\mathcal{J}_{m,n}(a, b, \alpha) \triangleq \sum_{r=0}^m \binom{m}{r} (-b)^{m-r} \left[\sum_{s=0}^{n+r} \frac{(n+r)^s b^{n+r-s}}{\alpha^{s+1} a^{m-s}} \text{Ei}(-b) - \frac{(n+r)^{n+r} e^{\alpha b/a}}{\alpha^{n+r+1} a^{m-n-r}} \text{Ei}\left(-\frac{\alpha b}{a} - b\right) + \frac{e^{-b}}{\alpha} \sum_{s=0}^{n+r-1} \sum_{u=0}^{n+r-s-1} \frac{u! (n+r)^s \binom{n+r-s-1}{u} b^{n+r-s-u-1}}{\alpha^s a^{m-s} (\alpha/a + 1)^{s+1}} \right]. \quad (20)$$

Furthermore, conditioned on $\left[\hat{\mathbf{G}}_u^\dagger[n-1]\right]_k$, $\hat{\mathbf{Y}}_i[n]$ is a zero-mean complex Gaussian vector with covariance matrix $\tilde{\mathbf{D}}_{li}$ which is independent of $\left[\hat{\mathbf{G}}_u^\dagger[n-1]\right]_k$. Thus, $\hat{\mathbf{Y}}_i[n] \sim \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{D}}_{li})$, independent of $\left[\hat{\mathbf{G}}_u^\dagger[n-1]\right]_k$. Hence, $\sum_{i=1}^L \left\|\hat{\mathbf{Y}}_i[n]\right\|^2$ is the sum of KL statistically independent but not necessarily identically distributed exponential RVs. According to [18, Theorem 2], we obtain

$$\sum_{i=1}^L \left\|\hat{\mathbf{Y}}_i[n]\right\|^2 \stackrel{\text{d}}{\sim} Y_k[n]. \quad (18)$$

Combining (16)–(18), we deduce (13). ■

Remark 1: In the general case, the PDF of the uplink SINR (13) accounts for both the effects of pilot contamination and Doppler shift. More specifically, the time variation of the channel decreases both the desired and interference signal powers by a factor of α^2 with comparison to the *zeroth* Doppler shift case, thus, degrading the SINR.

Remark 2: Increasing the relative velocity, the SINR presents ripples with peak and zero points following the behaviour of the $J_0(\cdot)$ Bessel function. In the marginal case of $\alpha = 1$, i.e., when there is no relative movement of the user, (13) expresses the downgrade of the system only to pilot contamination. Especially, if we assume very long training intervals (orthogonal pilot sequences) and no time variation, which is not practical in common scenarios with large number of antennas and moving users, our result coincides with [13, Eq. (6)]. At the other end, high velocity meaning $\alpha \rightarrow 0$ leads to zero SINR.

Corollary 1: Consider the high uplink power regime. We have

$$\gamma_k \stackrel{\text{d}}{\sim} \frac{\alpha^2 X_k[n-1]}{\alpha^2 C X_k[n-1] + Y_k[n]}, \quad \text{as } p_r \rightarrow \infty. \quad (19)$$

This corollary brings an important insight on the system performance, when p_r increases

$$\begin{aligned} \mathcal{I}_1 \triangleq & \sum_{t=0}^{N-K} \left[-e^{\frac{1}{\hat{\beta}_{lk}\alpha^2 p_r (C+1)}} \mathcal{J}_{q-1, N-K-t} \left(\frac{1}{\hat{\beta}_{lk}\alpha^2 (C+1)}, \frac{1}{\hat{\beta}_{lk}\alpha^2 p_r (C+1)}, \frac{1}{\mu_{k,p}} - \frac{1}{\hat{\beta}_{lk}\alpha^2 (C+1)} \right) \right. \\ & \left. + \sum_{u=1}^{N-K-t} \frac{(u-1)! (-1)^u p_r^{-q}}{\left(\hat{\beta}_{lk}\alpha^2 p_r (C+1) \right)^{N-K-t-u}} \Gamma(q) U \left(q, q+1+N-K-t-u, \frac{1}{\mu_{k,p} p_r} \right) \right] \quad (22) \end{aligned}$$

$$\begin{aligned} \mathcal{I}_2 \triangleq & \sum_{t=0}^{N-K} \left[-e^{\frac{1}{\hat{\beta}_{lk}\alpha^2 p_r C}} \mathcal{J}_{q-1, N-K-t} \left(\frac{1}{\hat{\beta}_{lk}\alpha^2 C}, \frac{1}{\hat{\beta}_{lk}\alpha^2 p_r C}, \frac{1}{\mu_{k,p}} - \frac{1}{\hat{\beta}_{lk}\alpha^2 C} \right) \right. \\ & \left. + \sum_{u=1}^{N-K-t} \frac{(u-1)! (-1)^u p_r^{-q}}{\left(\hat{\beta}_{lk}\alpha^2 p_r C \right)^{N-K-t-u}} \Gamma(q) U \left(q, q+1+N-K-t-u, \frac{1}{\mu_{k,p} p_r} \right) \right], \quad (23) \end{aligned}$$

asymptotically. As seen in (19), there is a finite SINR ceiling due to the simultaneous increment of the desired signal power as well as of the interference and channel estimation error powers.

Having obtained the PDF of the SINR, and by defining the function $\mathcal{J}_{m,n}(a, b, \alpha)$ as in (20) shown at the top of the previous page, where $\text{Ei}(\cdot)$ denotes the exponential integral function [19, Eq. (8.211.1)], we first obtain the exact $R_{lk}(p_r, \alpha)$ and a simpler lower bound $R_L(p_r, \alpha)$ as follows:

Theorem 1: The uplink ergodic achievable rate of transmission between the k th user in the l th cell to its BS for any finite number of antennas, under delayed channels, is given by

$$R_{lk}(p_r, \alpha) = \sum_{p=1}^{\varrho(\mathcal{A}_k) \tau_p(\mathcal{A}_k)} \sum_{q=1}^{\mathcal{X}_{p,q}(\mathcal{A}_k)} \frac{\mu_{k,p}^{-q} \log_2 e}{(q-1)!(N-K)! \hat{\beta}_{lk}^{N-K+1}} (\mathcal{I}_1 - \mathcal{I}_2), \quad (21)$$

where \mathcal{I}_1 and \mathcal{I}_2 are given by (22) and (23) shown at the top of next page, and where $U(\cdot, \cdot, \cdot)$ is the confluent hypergeometric function of the second kind [19, Eq. (9.210.2)].

Proof: See Appendix A. ■

In the case that all diagonal elements of \mathcal{A}_k are distinct, we have $\varrho(\mathcal{A}_k) = KL$, $\tau_p(\mathcal{A}_k) = 1$, and $\mathcal{X}_{p,1}(\mathcal{A}_k) = \prod_{q=1, q \neq p}^{KL} \left(1 - \frac{\mu_{k,q}}{\mu_{k,p}}\right)^{-1}$. The uplink rate becomes

$$R_{lk}(p_r, \alpha) = \sum_{p=1}^{KL} \sum_{t=0}^{N-K} \frac{\prod_{q=1, q \neq p}^{KL} \left(1 - \frac{\mu_{k,q}}{\mu_{k,p}}\right)^{-1} \log_2 e}{(N-K-t)! (-1)^{N-K-t} \mu_{k,p}} (\bar{\mathcal{I}}_1 - \bar{\mathcal{I}}_2), \quad (24)$$

where $\bar{\mathcal{I}}_1$ and $\bar{\mathcal{I}}_2$ are given by (25) and (26) shown at the top of next page. Note that, we have used the identity $U(1, b, c) = e^x x^{1-b} \Gamma(b-1, x)$ [29, Eq. (07.33.03.0014.01)] to obtain (24).

$$\begin{aligned} \bar{\mathcal{I}}_1 = & \sum_{t=0}^{N-K} \left[-e^{\frac{1}{\hat{\beta}_{lk}\alpha^2 p_r(C+1)}} \mathcal{J}_{0,N-K-t} \left(\frac{1}{\hat{\beta}_{lk}\alpha^2(C+1)}, \frac{1}{\hat{\beta}_{lk}\alpha^2 p_r(C+1)}, \frac{1}{\mu_{k,p}} - \frac{1}{\hat{\beta}_{lk}\alpha^2(C+1)} \right) \right. \\ & \left. + \sum_{u=1}^{N-K-t} \frac{(u-1)!(-1)^u}{\left(\hat{\beta}_{lk}\alpha^2(C+1)\right)^{N-K-t-u}} e^{\frac{1}{\mu_{k,p} p_r}} \mu_{k,p}^{N+1-K-t-u} \Gamma\left(N+1-K-t-u, \frac{1}{\mu_{k,p} p_r}\right) \right] \end{aligned} \quad (25)$$

$$\begin{aligned} \bar{\mathcal{I}}_2 = & \sum_{t=0}^{N-K} \left[-e^{\frac{1}{\hat{\beta}_{lk}\alpha^2 p_r C}} \mathcal{J}_{1,N-K-t} \left(\frac{1}{\hat{\beta}_{lk}\alpha^2 C}, \frac{1}{\hat{\beta}_{lk}\alpha^2 p_r C}, \frac{1}{\mu_{k,p}} - \frac{1}{\hat{\beta}_{lk}\alpha^2 C} \right) \right. \\ & \left. + \sum_{u=1}^{N-K-t} \frac{(u-1)!(-1)^u}{\left(\hat{\beta}_{lk}\alpha^2 C\right)^{N-K-t-u}} e^{\frac{1}{\mu_{k,p} p_r}} \mu_{k,p}^{N+1-K-t-u} \Gamma\left(N+1-K-t-u, \frac{1}{\mu_{k,p} p_r}\right) \right]. \end{aligned} \quad (26)$$

$$P_{\text{out}}(\gamma_{\text{th}}) = \begin{cases} 1, & \text{if } \gamma_{\text{th}} \geq 1/C \\ 1 - e^{-\frac{\gamma_{\text{th}}}{\hat{\beta}_{lk}(\alpha^2 p_r - \alpha^2 p_r C \gamma_{\text{th}})}} \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \sum_{t=0}^{N-K} \sum_{s=0}^t \binom{t}{s} \mathcal{X}_{p,q}(\mathcal{A}_k) \frac{\mu_{k,p}^{-q}}{(q-1)!} \Gamma(s+q) \left(\hat{\beta}_{lk}(\alpha^2 - \alpha^2 C \gamma_{\text{th}}) \right)^{s+q}, & \text{if } \gamma_{\text{th}} < 1/C. \end{cases} \quad (29)$$

Proposition 2: The uplink ergodic rate from the k th user in the l th cell to its BS, considering delayed channels, can be presented by a certain lower bound $R_L(p_r, \alpha)$:

$$\begin{aligned} R_{lk}(p_r, \alpha) & \geq R_L(p_r, \alpha) \\ & \triangleq \log_2 \left(1 + \frac{1}{C + \frac{1}{(N-K)\alpha^2 \hat{\beta}_{lk}} \left(\sum_{i=1}^L \sum_{k=1}^K (\beta_{lik} - \alpha^2 \hat{\beta}_{lik}) + \frac{1}{p_r} \right)} \right). \end{aligned} \quad (27)$$

Proof: See Appendix B. ■

1) *Outage Probability:* Bearing in mind that we investigate a block fading model, the study of the outage probability is of crucial interest. Basically, it defines the probability that the instantaneous SINR γ_k falls below a given threshold value γ_{th} :

$$P_{\text{out}}(\gamma_{\text{th}}) = \Pr(\gamma_k \leq \gamma_{\text{th}}). \quad (28)$$

Theorem 2: The outage probability of transmission from the k th user in the l th cell to its BS is given by (29), shown at the top of the page.

Proof: See Appendix C. ■

B. Characterization in the Low-SNR Regime

Even though Theorem 1 renders possible the exact derivation of the uplink sum-rate, it appears deficient to provide an insightful dependence on the various parameters such as the number of BS antennas and the transmit power. On that account, the study of the low power cornerstone, i.e., the low-SNR regime, is of great significance. There is no reason to consider the high-SNR regime, because in this regime an important metric such as the high-SNR slope $\mathcal{S}_\infty = \lim_{p_r \rightarrow 0} \frac{R_{lk}(p_r, \alpha)}{\log_2 p_r}$ [28] is zero due to the finite sum-rate, as shown in (19).

1) *Low-SNR Regime:* In case of low-SNR, it is possible to represent the rate by means of second-order Taylor approximation as

$$R_{lk}(p_r, \alpha) = \dot{R}_{lk}(0, \alpha) p_r + \ddot{R}_{lk}(0, \alpha) \frac{p_r^2}{2} + o(p_r^2), \quad (30)$$

where $\dot{R}_{lk}(p_r, \alpha)$ and $\ddot{R}_{lk}(p_r, \alpha)$ denote the first and second derivatives of $R_{lk}(p_r, \alpha)$ with respect to SNR p_r . In fact, these parameters enable us to examine the energy efficiency in the regime of low-SNR by means of two key element parameters, namely the minimum transmit energy per information bit, $\frac{E_b}{N_{0\min}}$, and the wideband slope S_0 [24]. Especially, we have

$$\frac{E_b}{N_{0\min}} = \lim_{p_r \rightarrow 0} \frac{p_r}{R_{lk}(p_r, \alpha)} = \frac{1}{\dot{R}_{lk}(0, \alpha)}, \quad (31)$$

$$S_0 = -\frac{2 \left[\dot{R}_{lk}(0, \alpha) \right]^2}{\ddot{R}_{lk}(0, \alpha)} \ln 2. \quad (32)$$

Theorem 3: In the low-SNR regime, the uplink sum-rate between the k th user in the l th cell to its BS in a multi-cell system, assuming delayed channels, can be captured by the minimum transmit energy per information bit, $\frac{E_b}{N_{0\min}}$, and the wideband slope S_0 , respectively, expressed by

$$\frac{E_b}{N_{0\min}} = \frac{\ln 2}{\alpha^2 (N - K + 1) \hat{\beta}_{lk}} \quad (33)$$

$$S_0 = \frac{-2(N - K + 1) / (N - K + 2)}{\alpha^4 + 2\alpha^2 C(N - K + 3) + \frac{2}{N - K + 2} \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k) \mu_{k,p}^{-q} q}{\hat{\beta}_{lk}(q-1)!}}. \quad (34)$$

Proof: See Appendix D. ■

C. Large Antenna Limit Analysis

In this section, we consider the large system limit by accounting for specific assumptions. Assuming constant transmit power p_r and : i) the number of BS antennas N grows infinitely large, while K is fixed, ii) both the number of users K and BS antennas N increase asymptotically by keeping their ratio $\kappa = \frac{N}{K}$ fixed; and with scaling the power with the number of antennas N , we obtain the corresponding SINRs, in order to scrutinize their properties. The purpose of this analysis is to exploit the reduction of the interference and thermal noise due to the property of orthogonal channels vectors between the BS and the users as $N \rightarrow \infty$ as well as to achieve increase of the sum-rate due to its dependence of N .

1) $N \rightarrow \infty$ with fixed p_r and K : Keeping in mind that an Erlang distribution with shape and scale parameters given by $N - K + 1$ and $\hat{\beta}_{lk}$, respectively, can be related with the sum of independent normal RVs $W_1[n-1], W_2[n-1], \dots, W_{2(N-K+1)}[n-1]$, $X_k[n-1]$ as follows:

$$X_k[n-1] = \frac{\hat{\beta}_{lk}}{2} \sum_{i=1}^{2(N-K+1)} W_i^2[n-1]. \quad (35)$$

By the substitution of (35) into (13) as well as by the use of the law of large numbers, the nominator and the first term of the denominator in (13) converge almost surely to $\alpha^2 p_r \hat{\beta}_{lk} / 2$ and $\alpha^2 p_r C \hat{\beta}_{lk} / 2$ as $N \rightarrow \infty$, while the second term of the denominator goes to 0. As a result, the deterministic equivalent of the SINR, $\bar{\gamma}_k$, when $N \rightarrow \infty$, is expressed as:

$$\gamma_k \xrightarrow{\text{a.s.}} \bar{\gamma}_k = \frac{1}{C}, \quad \text{as } N \rightarrow \infty. \quad (36)$$

The bounded SINR is expected because it is already known that as the number of BS antennas tends to infinity, both the intra-cell interference and noise are cancelled out, while the inter-cell interference due to pilot contamination remains.

2) $K, N \rightarrow \infty$ with fixed p_r and $\kappa = N/K$: In practice, the number of serving users K in each cell of next generation systems is not much less than the number of BS antennas N . In such case, the application of the law of numbers does not stand because the channel vectors between the BS and the users are not anymore pairwise orthogonal. This in turn induces new properties at the scenario under study, which are going to be revealed after the following analysis. Basically, we are going to derive the deterministic approximation $\bar{\gamma}_k$ of the SINR γ_k such that

$$\gamma_k - \bar{\gamma}_k \xrightarrow[N \rightarrow \infty]{\text{a.s.}} 0, \quad (37)$$

where $\xrightarrow[N \rightarrow \infty]{\text{a.s.}}$ denotes almost sure convergence.

Theorem 4: The deterministic equivalent $\bar{\gamma}_k$ of the uplink SINR between user k and its BS with ZF decoder is given by

$$\bar{\gamma}_k = \frac{\alpha^2 \hat{\beta}_{lk} (\kappa - 1)}{\alpha^2 C \hat{\beta}_{lk} (\kappa - 1) + \sum_{i=1}^L \frac{1}{K} \text{Tr} \tilde{\mathbf{D}}_{li}}. \quad (38)$$

Proof: Knowing that $\hat{\mathbf{Y}}_i[n] \sim \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{D}}_{li})$, we can write it as:

$$\hat{\mathbf{Y}}_i[n] = \mathbf{a}_i^H \tilde{\mathbf{D}}_{li}^{\frac{1}{2}}, \quad (39)$$

where $\mathbf{a}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_K)$. By substituting (35) as well as (39) into (13), we have

$$\gamma_k = \frac{\alpha^2 p_r \frac{\hat{\beta}_{lk}}{2} \sum_{i=1}^{2(N-K+1)} W_i^2[n-1]}{\alpha^2 p_r C \frac{\hat{\beta}_{lk}}{2} \sum_{i=1}^{2(N-K+1)} W_i^2[n-1] + p_r \sum_{i=1}^L \mathbf{a}_i^H \tilde{\mathbf{D}}_{li} \mathbf{a}_i + 1}. \quad (40)$$

Next, if we divide both the nominator and denominator of (40) by $2(N-K+1)$ and by using [10, Lemma 1], under the assumption that $\tilde{\mathbf{D}}_{li}$ has uniformly bounded spectral norm with respect to K , we arrive at the desired result (38). \blacksquare

Remark 3: Interestingly, in contrast to (36), the SINR is now affected by intra-cell interference as well as inter-cell interference and it is independent of the transmit power. In fact, the former justifies the latter, since both the desired and interference signals are changed by the same factor, if each user changes its power. Note that the interference terms remain because they depend on both N and K ; however, the dependence of thermal noise only from N makes it vanish. As expected, (38) coincides with (42), if $N \gg K$, i.e., when $\kappa \rightarrow \infty$, the SINR goes asymptotically to $1/C$.

Next, the deterministic equivalent sum-rate can be obtained by means of the dominated convergence [25] and the continuous mapping theorem [26] as

$$R_{lk}(p_r, \alpha) - \log_2(1 + \bar{\gamma}_k) \xrightarrow[N \rightarrow \infty]{\text{a.s.}} 0. \quad (41)$$

3) *Power-Scaling Law:* Let consider $p_r = E/\sqrt{N}$, where E is fixed regardless of N . Given that $\hat{\beta}_{lk}$ depends on $p_{\text{tr}} = \frac{\tau E}{\sqrt{N}}$, we have from (36) that for fixed K and $N \rightarrow \infty$,

$$\gamma_k \xrightarrow{\text{a.s.}} \frac{\alpha^2 \tau E^2 \beta_{lk}^2}{\alpha^2 \tau E^2 C \beta_{lk}^2 + 1}, \quad (42)$$

which is a non-zero constant. This implies that, we can reduce the transmit power proportionally to $1/\sqrt{N}$, while remaining a given quality-of-service. In the case where the BS has perfect CSI

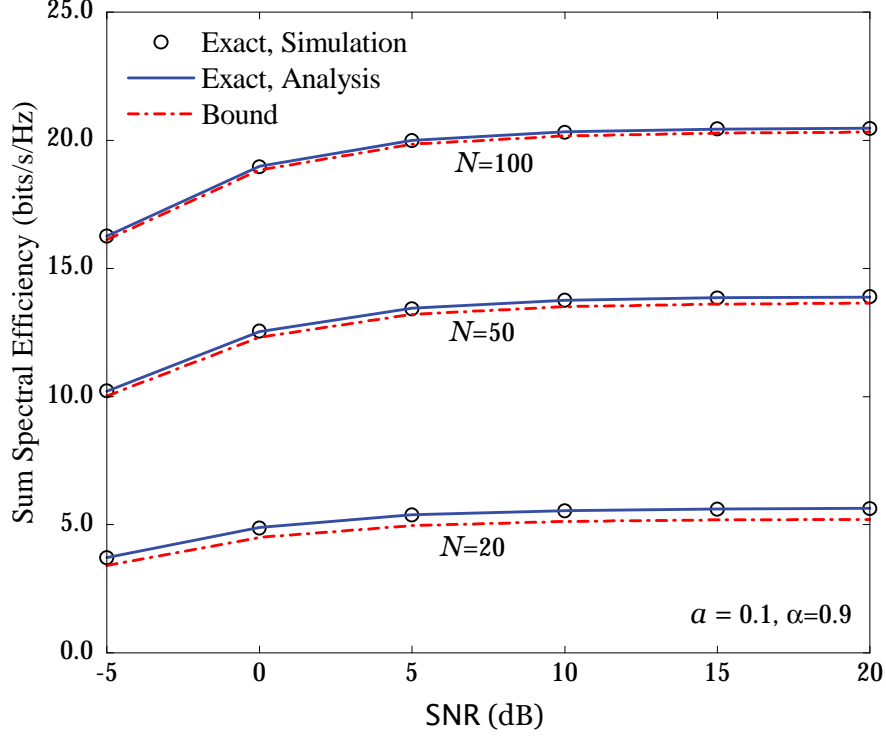


Fig. 1. Sum spectral efficiency versus SNR for different N ($a = 0.1$ and $\alpha = 0.9$).

and where there is no relative movement of the users, the result (42) is identical with the result in [13].

IV. NUMERICAL RESULTS

In this section, we present numerical results to verify our analysis by considering a cellular network with $L = 7$ cells and $K = 10$ users per cell. The coherence interval is $T = 200$ symbols (which corresponds to a coherence bandwidth of 200 kHz and a coherence time of 1 ms) and the length of training duration is $\tau = K$ symbols. Regarding the large-scale coefficients β_{lik} , we assume a simple scenario: $\beta_{ulk} = 1$ and $\beta_{lik} = a$, for $k = 1, \dots, K$, and $i \neq l$. Note that a can be considered as an intercell interference factor. In all examples, we choose $a = 0.1$. Furthermore, we define $\text{SNR} \triangleq p_r$.

In the following, we will examine the sum spectral efficiency which is defined as:

$$\mathcal{S}_l \triangleq \left(1 - \frac{\tau}{T}\right) \sum_{k=1}^K R_{lk}(p_r, \alpha), \quad (43)$$

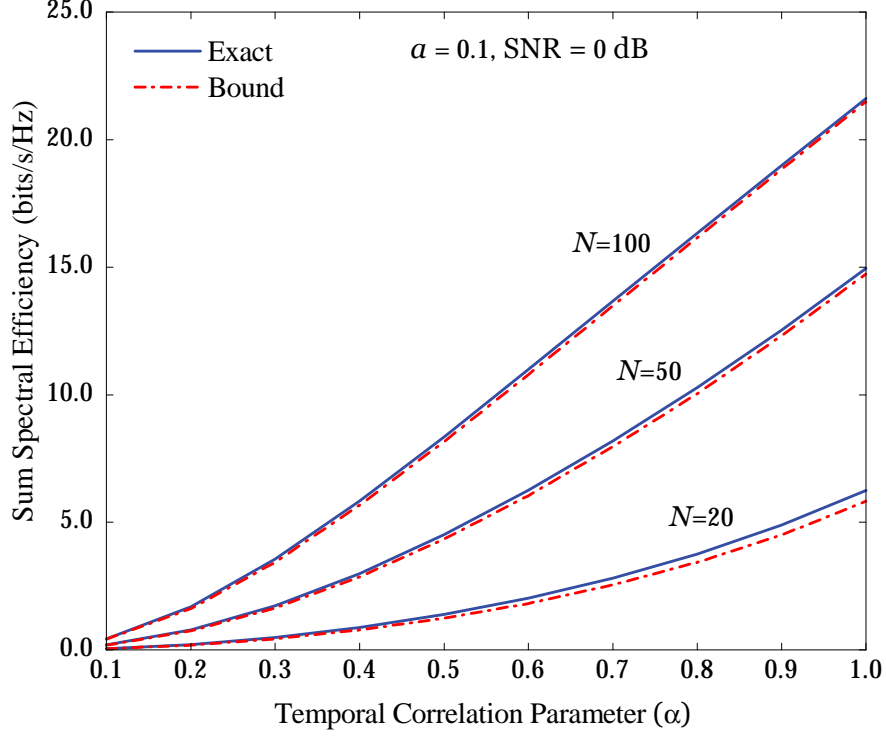


Fig. 2. Sum spectral efficiency versus α for different N ($a = 0.1$ and $\text{SNR} = 0 \text{ dB}$).

where $R_{lk}(p_r, \alpha)$ is given by (21).

Figure 1 shows the sum spectral efficiency versus the SNR for $N = 20, 50$, and 100 , at the intercell interference factor $a = 0.1$, and the temporal correlation parameter $\alpha = 0.9$. The “Exact, Analysis” curves are computed by using (21), the “Exact, Simulation” curves are generated via (12) using Monte-Carlo simulations, while the “Bound” curves are obtained by using the bound result given in Proposition 2. The exact match between the analytical and simulated results validates our analysis. It can be seen from the figure that, the proposed bound is very tight, especially for large antenna arrays. Furthermore, we can see that, at high SNR, the sum spectral efficiency saturates. This is due to the fact that when SNR increases, both the desired signal power and intercell interference power are increased. To improve the system performance, we can use more antennas at the BS. At $\text{SNR} = 5 \text{ dB}$, the sum spectral efficiencies can be increased by the factors of 2.5 or 5.5 if we increase N from 20 to 50 or from 20 to 100, respectively.

Next, we examine the effect of the temporal correlation parameter on the system performance as well as the tightness of our proposed bound given in Proposition 2. Figure 2 presents the

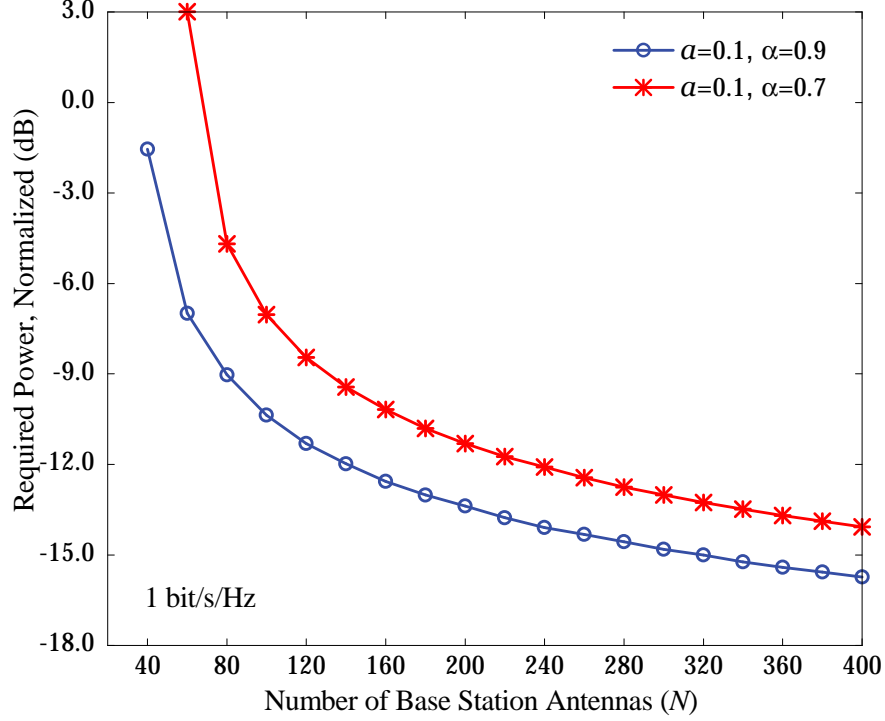


Fig. 3. Transmit power required to achieve 1 bit/s/Hz per user versus N ($a = 0.1$, $\alpha = 0.7$ and $\alpha = 0.9$).

sum spectral efficiency as a function of the temporal correlation parameter α , at $\text{SNR} = 0\text{dB}$, for $N = 20, 50$, and 100 . We can see that the system performance degrades significantly when the temporal correlation parameter decreases (or the time variation in the channel increases). Half of the spectral efficiency is reduced when α reduces from 1 to 0.6. Furthermore, at low α , using more antennas at the BS does not help improve the system performance much. Regarding the tightness of the proposed bound, we can see that the bound is very tight across the entire temporal correlation range.

Figure 3 shows the transmit power, p_r , that is needed to reach 1 bit/s/Hz per user. Here, we choose $a = 0.1$ and $\alpha = 0.7$ or 0.9 . As expected, the required transmit power decreases significantly when we increase the number of BS antennas. By doubling the number of BS antennas, we can cut back the transmit power by approximately 1.5dB. This observation is in line with the results of [6].

To further verify our analysis on large antenna limits, we consider Figure 4. Figure 4 shows the sum spectral efficiency versus the number of base station antennas for different values of α , and

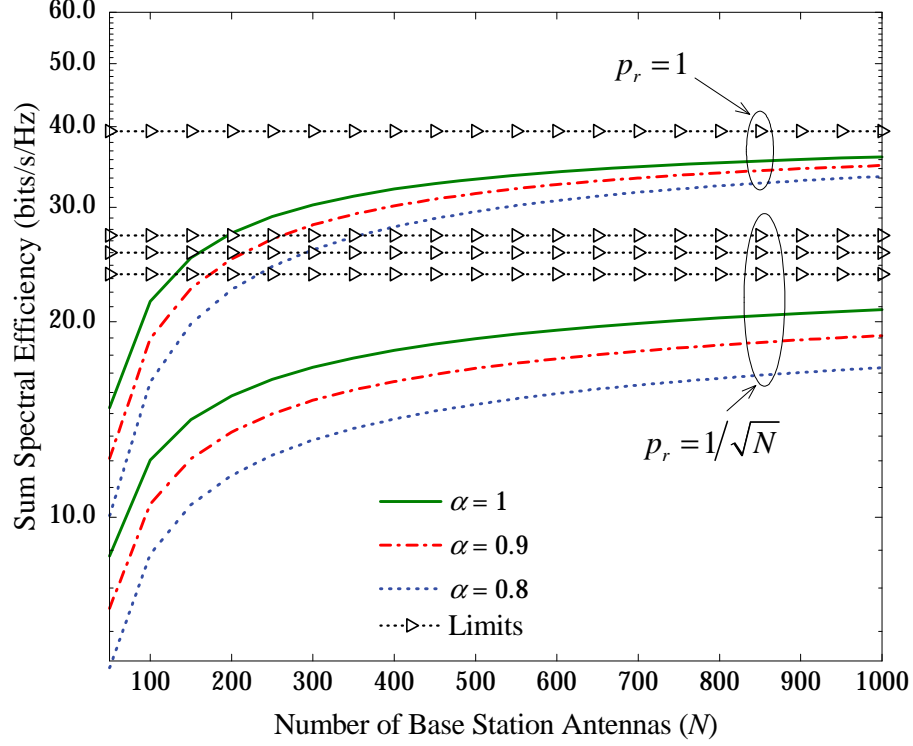


Fig. 4. Sum spectral efficiency versus N for different α .

for two cases: the transmit power, p_r , is fixed regardless of N , and the transmit power is scaled as $p_r = 1/\sqrt{N}$. The “Limits” curves are computed via the results obtained in Section III-C. As expected, as the number of the base station antennas increases, the sum spectral efficiencies converge to their limits. When the transmit power is fixed, the asymptotic performance (as $N \rightarrow \infty$) does not depend on the temporal correlation parameter. By contrast, when the transmit power is scaled as $1/\sqrt{N}$, the asymptotic performance depends on α .

Finally, we consider the outage performance versus SNR at $N = 100$, for different temporal correlation parameters ($\alpha = 1$, and 0.9), and for different threshold values ($\gamma_{\text{th}} = 2$, and 3). See Figure 5. We can see that, the outage probability strongly depends on α . At SNR = 0 dB, by reducing α from 1 to 0.9, the outage probability increases from 7×10^{-6} to 5×10^{-3} , and from 3×10^{-2} to 5×10^{-1} for $\gamma_{\text{th}} = 2$, and 3, respectively. In addition, the outage probability significantly improves when the threshold values are slightly reduced. This is due to the fact that, with large antenna arrays, the channel hardening occurs, and hence, the SINR concentrates around its mean. As a results, by slightly reducing the threshold values, we can obtain a very

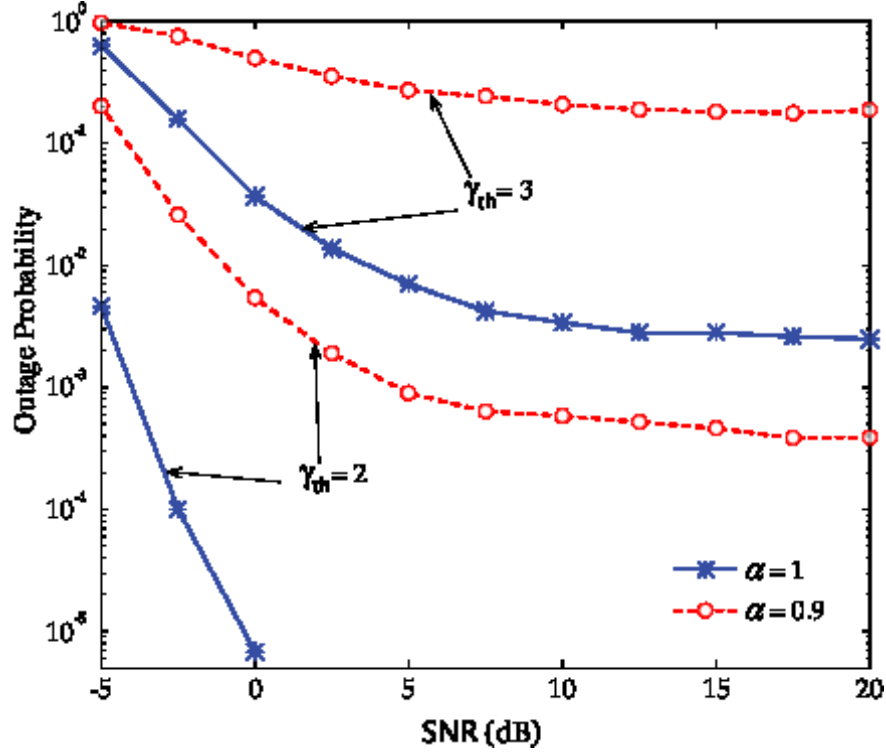


Fig. 5. Outage probability versus SNR for different α and γ_{th} ($N = 100$).

low outage probability.

V. CONCLUSIONS

In this paper, we characterized the uplink performance of a cellular network taking into account both the well-known pilot contamination and the unavoidable, but less studied, time variation. The latter effect, inherent in the vast majority of propagation scenarios, stems from the user mobility. Summarizing the main contributions of this work, new analytical closed-form expressions for the PDF of the SINR and the corresponding achievable sum-rate, that hold for any finite number of antennas, were derived. Moreover, a complete investigation of the low-SNR regime took place. Nevertheless, asymptotic expressions in the large antenna/user limit were also obtained, as well as the power-scaling law was studied. As a final point, numerical illustrations depicted how the time variation affects the performance in various Doppler shifts for finite and infinite number of antennas. Notably, the outcome is that large number of antennas should be preferred even in time-varying conditions.

APPENDIX

A. Proof of Theorem 1

The uplink ergodic rate from the k th user in the l th cell to its BS (in *bits/s/Hz*) is given by

$$\begin{aligned}
R_{lk}(p_r, \alpha) &= \mathbb{E}_{X_k, Y_k} \left\{ \log_2 \left(1 + \frac{p_r \alpha^2 X_k[n-1]}{p_r \alpha^2 C X_k[n-1] + p_r Y_k[n] + 1} \right) \right\} \\
&= \int_0^\infty \int_0^\infty \log_2 \left(1 + \frac{p_r \alpha^2 x}{p_r \alpha^2 C x + p_r y + 1} \right) p_{X_k}(x) p_{Y_k}(y) dx dy \\
&= \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k) \mu_{k,p}^{-q} \log_2 e}{(q-1)! (N-K)! \hat{\beta}_{llk}^{N-K+1}} \\
&\quad \times \int_0^\infty \int_0^\infty \ln \left(1 + \frac{p_r \alpha^2 x}{p_r \alpha^2 C x + p_r y + 1} \right) x^{N-K} e^{\frac{-x}{\hat{\beta}_{llk}}} y^{q-1} e^{\frac{-y}{\mu_{k,p}}} dx dy \\
&= \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k) \mu_{k,p}^{-q} \log_2 e}{(q-1)! (N-K)! \hat{\beta}_{llk}^{N-K+1}} \\
&\quad \times \underbrace{\left(\int_0^\infty \int_0^\infty \ln \left(1 + \frac{p_r \alpha^2 (C+1) x}{p_r y + 1} \right) x^{N-K} e^{\frac{-x}{\hat{\beta}_{llk}}} y^{q-1} e^{\frac{-y}{\mu_{k,p}}} dx dy \right)}_{\triangleq \mathcal{I}_1} \\
&\quad - \underbrace{\int_0^\infty \int_0^\infty \ln \left(1 + \frac{p_r \alpha^2 C x}{p_r y + 1} \right) x^{N-K} e^{\frac{-x}{\hat{\beta}_{llk}}} y^{q-1} e^{\frac{-y}{\mu_{k,p}}} dx dy}_{\triangleq \mathcal{I}_2} \tag{44}
\end{aligned}$$

$$= \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \frac{\mathcal{X}_{p,q}(\mathcal{A}_k) \mu_{k,p}^{-q} \log_2 e}{(q-1)! (N-K)! \hat{\beta}_{llk}^{N-K+1}} (\mathcal{I}_1 - \mathcal{I}_2). \tag{45}$$

We first derive \mathcal{I}_1 by evaluating the integral over x . By using [19, Eq. (4.337.5)], we obtain

$$\begin{aligned}
\mathcal{I}_1 &= \sum_{t=0}^{N-K} \int_0^\infty \left[-f(y)^{N-K-t} e^{-f(y)} \text{Ei}(f(y)) \right. \\
&\quad \left. + \sum_{u=1}^{N-K-t} (u-1)! f(y)^{N-K-t-u} \right] y^{q-1} e^{\frac{-y}{\mu_{k,p}}} dy, \tag{46}
\end{aligned}$$

where $f(y) \triangleq -\frac{p_r y + 1}{\hat{\beta}_{llk} p_r \alpha^2 (C+1)}$. Using [23, Lemma 1] and [21, Eq. (39)], we can easily obtain \mathcal{I}_1 as given in (22). Similarly, we obtain \mathcal{I}_2 as given in (23). Substitution of \mathcal{I}_1 and \mathcal{I}_2 into (45) concludes the proof.

B. Proof of Proposition 2

By using Jensen's inequality, we have

$$\begin{aligned} R_{lk}(p_r, \alpha) &= \mathbb{E} \{ \log_2 (1 + \gamma_k) \} = \mathbb{E} \left\{ \log_2 \left(1 + \frac{1}{1/\gamma_k} \right) \right\} \\ &\geq \log_2 \left(1 + \frac{1}{\mathbb{E} \{ 1/\gamma_k \}} \right) \triangleq R_L(p_r, \alpha). \end{aligned} \quad (47)$$

To compute $R_L(p_r, \alpha)$, we need to compute $\mathbb{E} \{ 1/\gamma_k \}$. From (16), we have

$$\begin{aligned} \mathbb{E} \left\{ \frac{1}{\gamma_k} \right\} &= C + \frac{1}{\alpha^2} \sum_{i=1}^L \mathbb{E} \left\{ \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \tilde{\mathbf{E}}_{li}[n] \right\|^2 \right\} \\ &\quad + \frac{1}{\alpha^2 p_r} \mathbb{E} \left\{ \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^2 \right\} \\ &= C + \frac{1}{\alpha^2} \mathbb{E} \left\{ \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^2 \right\} \left(\sum_{i=1}^L \sum_{k=1}^K (\beta_{lik} - \alpha^2 \hat{\beta}_{lik}) + \frac{1}{p_r} \right) \\ &= C + \frac{1}{(N-K)\alpha^2 \hat{\beta}_{ulk}} \left(\sum_{i=1}^L \sum_{k=1}^K (\beta_{lik} - \alpha^2 \hat{\beta}_{lik}) + \frac{1}{p_r} \right). \end{aligned} \quad (48)$$

In the third equality of (48), we have considered the independence between the two variables, while in the last equality, we have used the following result:

$$\begin{aligned} \mathbb{E} \left\{ \left\| [\hat{\mathbf{G}}_u^\dagger[n-1]]_k \right\|^2 \right\} &= \mathbb{E}_{X_k} \left\{ \frac{1}{X_k[n-1]} \right\} \\ &= \int_0^\infty \frac{e^{-x/\hat{\beta}_{ulk}}}{(N-K)! \hat{\beta}_{ulk}^2} \left(\frac{x}{\hat{\beta}_{ulk}} \right)^{N-K-1} dx = \frac{1}{(N-K) \hat{\beta}_{ulk}}. \end{aligned} \quad (49)$$

Note that we have used [19, Eq. (3.326.2)] to obtain (49). Thus, the desired result (27) is obtained from (47) and (48).

C. Proof of Theorem 2

Clearly, from (13), $\gamma_k < 1/C$. Thus, if $\gamma_{\text{th}} \geq 1/C$, then $P_{\text{out}}(\gamma_{\text{th}}) = 1$. Next, we consider the case where $\gamma_{\text{th}} < 1/C$. Taking the probability of the instantaneous SINR γ_k , given by (13), we

$$\ddot{R}_{lk}(p_r, \alpha) = \frac{1}{\ln 2} \mathbb{E}_{X_k, Y_k} \left\{ \frac{\alpha^2 X_k [n-1] \left(\alpha^4 X_k^2 [n-1] + 2\varsigma_k (1 + p_r \varsigma_k)^2 (1 + \alpha^2 p_r X_k [n-1] + p_r \varsigma_k) \right)}{(\alpha^4 p_r^2 (C+1) X_k^2 [n-1] + p Y_k [n] + 1)^2 (\alpha^4 p_r^2 C X_k^2 [n-1] + p_r Y_k [n] + 1)^4} \right\}, \quad (56)$$

can determine the outage probability as

$$\begin{aligned} P_{\text{out}} &= \Pr \left(\frac{\alpha^2 p_r X_k}{\alpha^2 p_r C X_k + p_r Y_k + 1} \leq \gamma_{\text{th}} \right) \\ &= \int_0^\infty \Pr \left(X_k < \frac{\gamma_{\text{th}} (p_r Y_k + 1)}{\alpha^2 p_r - \gamma_{\text{th}} \alpha^2 p_r C} \mid Y_k \right) p_{Y_k}(y) dy \\ &= 1 - e^{-\frac{\gamma_{\text{th}}}{p_r \gamma_{\text{th}}}} \sum_{t=0}^{N-K} \int_0^\infty e^{\frac{-y}{\gamma_{\text{th}}}} \sum_{t=0}^{N-K} \frac{\left(\frac{\gamma_{\text{th}}}{\gamma_{\text{th}}} \right)^t}{t!} \left(y + \frac{1}{p_r} \right)^t p_{Y_k}(y) dy \\ &= 1 - e^{-\frac{\gamma_{\text{th}}}{p_r \gamma_{\text{th}}}} \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \sum_{t=0}^{N-K} \mathcal{X}_{p,q}(\mathcal{A}_k) \frac{\mu_{k,p}^{-q}}{(q-1)!} \frac{\left(\frac{\gamma_{\text{th}}}{\gamma_{\text{th}}} \right)^t}{t!} \\ &\quad \times \int_0^\infty y^{q-1} e^{\frac{-y}{\gamma_{\text{th}}}} \left(y + \frac{1}{p_r} \right)^t dy \\ &= 1 - e^{-\frac{\gamma_{\text{th}}}{p_r \gamma_{\text{th}}}} \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \sum_{t=0}^{N-K} \sum_{s=0}^t \binom{t}{s} \mathcal{X}_{p,q}(\mathcal{A}_k) \frac{\Gamma(s+q) \bar{\gamma}_{\text{th}}^{s+q}}{\mu_{k,p}^q (q-1)!}, \end{aligned} \quad (50)$$

where $\bar{\gamma}_{\text{th}} \triangleq \hat{\beta}_{llk} (\alpha^2 - \alpha^2 C \gamma_{\text{th}})$, and where in the third equality, we have used that the cumulative density function of X_k (Erlang variable) is

$$\begin{aligned} F_{X_k}(x) &= \Pr(X_k \leq x) \\ &= 1 - \exp \left(-\frac{x}{\hat{\beta}_{llk}} \right) \sum_{t=0}^{N-K} \frac{1}{t!} \left(\frac{x}{\hat{\beta}_{llk}} \right)^t. \end{aligned} \quad (51)$$

The last equality of (50) was derived after applying the binomial expansion of $(y + 1/p_r)^t$ and [19, Eq. (3.351.1)].

D. Proof of Theorem 3

The initial step for the derivation of the minimum transmit energy per information bit is to cover the need for exact expressions regarding the derivatives of $R_{lk}(p_r, \alpha)$. In particular, this

can be given by

$$\begin{aligned} \dot{R}_{lk}(p_r, \alpha) &= \frac{1}{\ln 2} \\ &\times \mathbb{E}_{X_k, Y_k} \left\{ \frac{\alpha^2 X_k[n-1] / (\alpha^4 p_r^2 C X_k^2[n-1] + p_r Y_k[n] + 1)}{(\alpha^4 p_r^2 (C+1) X_k^2[n-1] + p_r Y_k[n] + 1)} \right\}. \end{aligned} \quad (52)$$

Easily, its value at $p_r = 0$ is

$$\dot{R}_{lk}(0, \alpha) = \frac{1}{\ln 2} \mathbb{E}_{X_k} \{ \alpha^2 X_k[n-1] \}. \quad (53)$$

Taking into account that $X_k[n-1]$ is Erlang distributed, its expectation can be written as

$$\mathbb{E}_{X_k} \{ X_k[n-1] \} = (N - K + 1) \hat{\beta}_{lk}. \quad (54)$$

Substituting (54) and (53) into (31), we obtain the desired result.

The second derivative of $R_{lk}(p_r, \alpha)$, needed for the evaluation of the wideband slope, is given by (56) shown at the top of the previous page, where $\varsigma_k \triangleq \alpha^2 C X_k[n-1] + Y_k[n]$. Hence, $\ddot{R}_{lk}(0, \alpha)$ can be expressed by

$$\begin{aligned} \ddot{R}_{lk}(0, \alpha) &= \frac{1}{\ln 2} \mathbb{E}_{X_k, Y_k} \{ \alpha^6 X_k^3[n-1] + 2\alpha^4 C X_k^2[n-1] \\ &\quad + 2\alpha^2 X_k[n-1] Y_k[n] \}. \end{aligned} \quad (57)$$

The moments of $X_k[n-1]$ are obtained by means of the corresponding derivatives of its moment generating function (MGF) at zero $M_{X_k}^{(n)}(0)$, i.e., $\mathbb{E}_{X_k} \{ X_k^n[n-1] \} = M_{X_k}^{(n)}(0)$. Thus, having in mind that the MGF of the Erlang distribution is

$$M_{X_k}(t) = \frac{1}{(1 - \hat{\beta}_{lk} t)^{N-K+1}}, \quad (58)$$

we can obtain the required moments of $X_k[n-1]$ as

$$\begin{aligned} \mathbb{E}_{X_k} \{ X_k^2[n-1] \} &= M_{X_k}^{(2)}(0) \\ &= \frac{\Gamma(N - K + 3)}{\Gamma(N - K + 1)} \hat{\beta}_{lk}^2 \end{aligned} \quad (59)$$

$$\begin{aligned} \mathbb{E}_{X_k} \{ X_k^3[n-1] \} &= M_{X_k}^{(3)}(0) \\ &= \frac{\Gamma(N - K + 4)}{\Gamma(N - K + 1)} \hat{\beta}_{lk}^3. \end{aligned} \quad (60)$$

In addition, since $X_k[n-1]$ and $Y_k[n]$ are uncorrelated, we have $\mathbb{E}_{X_k, Y_k} \{X_k[n-1]Y_k[n]\} = \mathbb{E}_{X_k} \{X_k[n-1]\} \mathbb{E}_{Y_k} \{Y_k[n]\}$. In other words, it is necessary to find the expectation of $Y_k[n]$. As aforementioned, the PDF of $Y_k[n]$ obeys (15) and has expectation given by definition as

$$\begin{aligned} \mathbb{E}_{Y_k} \{Y_k[n]\} &= \int_0^\infty y p_{Y_k}(y) dy \\ &= \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \mathcal{X}_{p,q}(\mathcal{A}_k) \frac{\mu_{k,p}^{-q}}{(q-1)!} \int_0^\infty y^q e^{\frac{-y}{\mu_{k,p}}} dy \\ &= \sum_{p=1}^{\varrho(\mathcal{A}_k)} \sum_{q=1}^{\tau_p(\mathcal{A}_k)} \mathcal{X}_{p,q}(\mathcal{A}_k) \frac{\mu_{k,p}^{-q} q}{(q-1)!}, \end{aligned} \quad (61)$$

where we have used [19, Eq. (3.326.2)] as well as the identity $\Gamma(q+1) = q!$. As a result, $\ddot{R}_{lk}(0, \alpha)$ follows by means of (59), (60), (61). Finally, substitution of the (53) and (57) into (32) yields the wideband slope.

REFERENCES

- [1] A. Papazafeiropoulos, H. Q. Ngo, M. Matthaiou, and T. Ratnarajah, "Uplink performance of conventional and massive MIMO cellular systems with delayed CSIT," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Washington, D.C., Sep. 2014.
- [2] "5G: A Technology Vision," Huawei Technologies Co., Ltd., Shenzhen, China, Whitepaper, Nov. 2013. [Online]. www.huawei.com/ilink/en/download/HW_314849
- [3] D. Gesbert, M. Kountouris, R. W. Heath Jr., C. B. Chae, and T. Sälzer, "Shifting the MIMO paradigm," *IEEE Sig. Proc. Mag.*, vol. 24, no. 5, pp. 36–46, Oct. 2007.
- [4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [5] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [7] C. J. Chen and L. C. Wang, "Performance analysis of scheduling in multiuser MIMO systems with zero-forcing receivers," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1435–1445, Sep. 2007.
- [8] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2006.
- [9] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [10] K. T. Truong and R. W. Heath, Jr., "Effects of channel aging in Massive MIMO Systems," *IEEE/KICS J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

- [11] A. Papazafeiropoulos and T. Ratnarajah “Linear precoding for downlink massive MIMO with delayed CSIT and channel prediction” in *Proc. IEEE WCNC 2014*, Apr. 2014, pp. 821–826.
- [12] A. Papazafeiropoulos and T. Ratnarajah “Uplink performance of massive MIMO subject to delayed CSIT and anticipated channel prediction,” in *Proc. IEEE ICASSP*, May 2014.
- [13] H. Q. Ngo, M. Matthaiou, T. Q. Duong, and E. G. Larsson, “Uplink performance analysis of multicell MU-SIMO systems with ZF receivers,” *IEEE Trans. Veh. Tech.*, vol. 62, no. 9, pp. 4471–4483, Nov. 2013.
- [14] S. Verdú, *Multiuser Detection*. Cambridge, UK: Cambridge University Press, 1998.
- [15] H. Q. Ngo, M. Matthaiou, and E. G. Larsson. “Performance analysis of large scale MU-MIMO with optimal linear receivers,” in *Proc. IEEE Swe-CTW*, Oct. 2012, pp. 59–64.
- [16] H. Shin and M. Z. Win, “MIMO diversity in the presence of double scattering,” *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2976–2996, Jul. 2008.
- [17] D. A. Gore, R. W. Heath Jr., and A. J. Paulraj, “Transmit selection in spatial multiplexing systems,” *IEEE Commun. Lett.*, vol. 6, no. 11, pp. 491–493, Nov. 2002.
- [18] A. Bletsas, H. Shin, and M. Z. Win, “Cooperative communications with outage-optimal opportunistic relaying,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3450–3460, Sep. 2007.
- [19] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. San Diego, CA: Academic, 2007.
- [20] Wolfram, “The Wolfram functions site.” Available: <http://functions.wolfram.com>
- [21] M. Kang and M.-S. Alouini, “Capacity of MIMO Rician channels,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 112–122, Jan. 2006.
- [22] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series, Volume 3: More Special Functions*. New York: Gordon and Breach Science, 1990.
- [23] H. Q. Ngo, T. Q. Duong, and E. G. Larsson, “Uplink performance analysis of multicell MU-MIMO with zero-forcing receivers and perfect CSI,” in *Proc. IEEE Swe-CTW*, Sweden, Nov. 2011, pp. 40–45.
- [24] S. Verdú, “Spectral efficiency in the wideband regime,” *IEEE Trans. Info. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
- [25] P. Billingsley, *Probability and Measure*, 3rd ed. John Wiley & Sons, Inc., 1995.
- [26] A. W. van der Vaart, *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, New York, 2000.
- [27] S. Shamai (Shitz) and S. Verdú, “The impact of frequency-flat fading on the spectral efficiency of CDMA,” *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.
- [28] A. Lozano, A. M. Tulino, and S. Verdú, “High-SNR power offset in multiantenna communications,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
- [29] Wolfram, “The Wolfram functions site.” Available: <http://functions.wolfram.com>